Scientia Pharmaceutica www.scipharm.at

Research article Open Access

QSAR Study on 6-Substituted Benzimidazoles: An Insight into the Structural Requirement for the Angiotensin II AT1 Receptor Antagonist

Anurekha JAIN * [1], Subhash C. CHATURVEDI [2]

[1] B. R. Nahata College of Pharmacy, Rajiv Gandhi Technical University, Mhow-Neemuch Road 458001 Mandasur (M.P.), India.
[2] School of Pharmacy, Devi Ahilya VishawaVidalaya, Takshashila Parser Khandawa Road, 452003 Indore (M.P.), India.

* Corresponding author. E-mail: anurekha_jain@yahoo.com (A. Jain)

Sci Pharm. 2009; 77: 555–565 doi:10.3797/scipharm.0901-30

Published: June 6th 2009 Received: January 30th 2009
Accepted: June 5th 2009

This article is available from: http://dx.doi.org/10.3797/scipharm.0901-30

Abstract

With an aim to identify the structural requirements for selective AT1 angiotensin antagonistic activity, a quantitative structure activity relationship (QSAR) analysis was carried out on a series of 6-substituted benzimidazole derivatives. The QSAR expressions were generated using 28 compounds and the predictive ability of the resulting model was evaluated against a test set of 12 compounds. The internal (cross validated squared correlation coefficient) and external consistency (predictive correlation coefficient) of the QSAR model was 0.78 and 0.40 respectively. In the present work QSAR analysis reveals that geometrical, structural, and shape descriptors govern the angiotensin II AT1 antagonistic activity.

Keywords

Nonpeptide Angiotensin II Antagonist • AT1 receptor • QSAR • Hypertension • RAS

Introduction

The rennin-angiotensin system (RAS) plays a major role in the regulation of blood pressure and electrolyte homeostasis [1]. RAS is a cascade of proteolytic enzymes (renin and angiotensin converting enzyme (ACE)) that result in the production of the systemic hormone angiotensin II (AII).The blockade of RAS with inhibitors of ACE has demonstrated the effectiveness of the reduction of levels of AII on cardiovascular and kidney

heamodynamics, aldosterone production and release, and the absorption of sodium. Antagonists of AII constitute an alternative method blocking the RAS. Several peptidic and nonpeptidic AII receptor antagonists are known. The therapeutic availability is less for the peptidic AII antagonist due to their poor bioavailability; short plasma half-life and partial agonist activity but the nonpeptidic AII receptors antagonist lack the defect of peptidic antagonist [2].The therapeutic profile of AII receptor antagonist is thought to be similar to that of angiotensin converting enzyme (ACE) inhibitors such as captopril, enalapril, and lisinopril. In addition, since AII receptor antagonist does not affect the metabolism of bradykinin so they may not have the side effect of ACE inhibitors, such as dry cough and angiodema. Recently, the QSAR analysis is a highly interested area for designing the compound before synthesis [3–5].

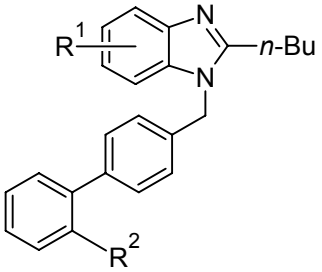
To gain insight into the structural and molecular requirement influencing the A II antagonistic activity, herein we depict QSAR analysis of some 6-substituted benzimidazoles derivative for AII antagonistic activity. The relevance of the model used for the design of novel derivatives should be assessed not only in terms of predictivity, either internal or external, but also in terms of their ability to provide a chemical and structural explanation of their binding interaction. These results should provide guidelines for design of more potent and selective AII antagonist.

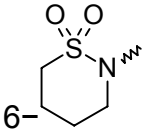
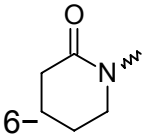
Experimental

The AII receptor antagonistic activity data of 6-substituted benzimidazole derivatives were taken from the reported work of Uwe Ries *et al.* [6] (Table 1). The biological activity data (IC_{50} in 10^{-7} M) was converted to negative logarithmic mole dose (pIC_{50}) for quantitative structure activity relationship (QSAR) analysis. The molecular modeling study was performed using CS ChemOffice [7] version 10 and DRAGON [8] program while the regression analysis was carried out with VALSTAT [9].

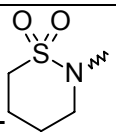
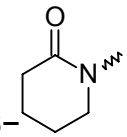
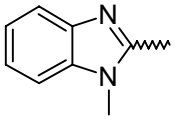
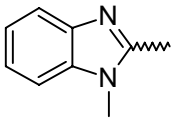
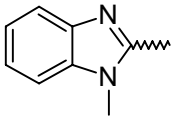
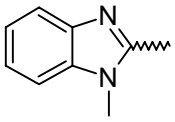
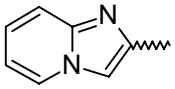
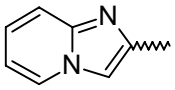
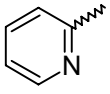
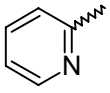
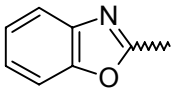
The compound Series was divided into training set of 28 compounds and test set of 12 compounds on the basis of structural diversity and with the aim that it should cover the complete range of variation in antagonist activity. The molecular structures of compounds were sketched by using Chem Draw [7] and then ChemUltra [7] used to convert them into 3D structures. The energy minimization of the molecule was done using molecular mechanics (MM2) until the RMS gradient value became smaller than 0.1 kcal/mol Å. The energy minimized molecules were subjected to the re-optimization via Austin model-1 (AM1) Hamiltonian method until the RMS gradient attained a value smaller than 0.0001 kcal/mol Å using MOPAC. The molecule was saved as MOL file format. Pursuly, the MOL file was used for the calculation of various Descriptors using DRAGON program. The data was used in order to establish a correlation between physicochemical parameters as independent variables and pIC_{50} as dependent variable employing sequential multiple linear regression analysis method by statistical program VALSTAT. In this regression analysis, the program searches all the permutations and combinations sequentially for the data set. The ± data within the parentheses are the standard deviation, associated with coefficient of descriptors in regression equations. The best model was selected from the various statistically significant equations on the basis of observed squared correlation coefficient (r^2), standard error of estimate (SEE), sequential Fischer test (F), bootstrapping squared correlation coefficient (r^2_{bs}), bootstrapping standard deviation (S_{bs}), cross validated squared correlation coefficient using leave one out procedure (Q^2), chance

Tab. 1. Structure, activities and Descriptors of 6-substituted Benzimidazoles used in training and test set



S. No.	R^1	R^2	IC_{50}[a]	pIC_{50}[b]	Mor21u[c]	Mor31u	Mor14m	MR[d]
1	H	COOH	400	6.39	−1.713	0.494	−0.54	114.416
2	4-CH_3	COOH	1200	5.92	−1.887	0.827	−0.543	119.457
3	5−CH_3	COOH	1200	5.92	−2.076	0.572	−0.487	119.457
4	6−CH_3	COOH	850	6.07	−2.076	0.572	−0.487	119.457
5	7−CH_3	COOH	480	6.31	−2.099	0.865	−0.562	119.457
6	4−NH_2	COOH	1700	5.76	−1.763	0.374	−0.766	119.116
7	5−NH_2	COOH	820	6.08	−1.426	0.473	−0.800	119.116
8	6−NH_2	COOH	540	6.26	−1.537	0.439	−0.866	119.116
9	7−NH_2	COOH	1060	5.97	−1.575	0.339	−0.367	119.116
10	4−(acetylamino)	COOH	5700	5.24	−2.292	0.456	−0.551	127.498
11	5−(acetylamino)	COOH	460	6.33	−1.682	0.506	−0.916	127.498
12	6−(acetylamino)	COOH	180	6.74	−1.662	0.654	−0.882	127.498
13	7−(acetylamino)	COOH	1800	5.74	−2.074	0.360	−0.699	127.498
14	4−$NHCONHC_6H_{11}$	COOH	29300	4.53	−2.425	0.204	−1.306	150.849
15	5−$NHCONHC_6H_{11}$	COOH	800	6.09	−2.072	0.152	−1.362	150.849
16	6−$NHCONHC_6H_{11}$	COOH	26	7.58	−2.253	0.178	−1.162	150.849
17	7−$NHCONHC_6H_{11}$	COOH	160	6.79	−1.976	0.436	−1.250	150.849
18	6−$CH_3(CH_2)_4NH$	COOH	390	6.40	−2.356	0.602	−0.453	143.084
19	6−(piperidin−1−yl)	COOH	160	6.79	−2.496	0.556	−0.366	136.107
20	6−(*n*-Bu−CONH)	COOH	86	7.06	−2.472	0.845	−0.832	141.327
21	6−$(CH_3)_2NCONH$	COOH	24	7.61	−2.095	0.035	−0.925	135.968
22	6−$C_6H_{11}NHCONCH_3$	COOH	26	7.58	−2.381	0.552	−1.205	150.339
23	6−[methyl(propyl−sulfonyl)amino]	COOH	33	7.48	−2.212	0.859	−0.757	143.139
24	 6−	COOH	34	7.46	−2.279	0.441	−0.632	140.148
25	 6−	COOH	81	7.09	−2.559	0.882	−0.701	138.631
26	6−$C_6H_{11}NHCONH$	Tetrazol−5−yl	21	7.67	−1.833	−0.976	0.217	161.485
27	6−$C_6H_{11}NHCONCH_3$	Tetrazol−5−yl	10	8	−2.764	0.703	−1.514	166.381
28	6−$(CH_3)_2NCONH$	Tetrazol−5−yl	8	8.09	−2.343	0.902	−0.98	143.246

Tab. 1. (Cont.)

S. No.	R^1	R^2	IC_{50}[a]	pIC_{50}[b]	Mor21u[c]	Mor31u	Mor14m	MR[d]
29	 6−	Tetrazol−5−yl	3	8.52	−2.589	0.829	−0.570	150.784
30	 6−	Tetrazol−5−yl	4	8.39	−2.836	1.036	−0.295	149.267
31	4−CH_3,  6−	COOH	3	8.52	−1.999	0.560	−0.819	153.05
32	 6−	COOH	3	8.52	−1.464	0.706	−1.137	148.009
33	4−CH_3,  6−	Tetrazol−5−yl	13	7.88	−2.403	0.322	−1.014	163.686
34	 6−	Tetrazol−5−yl	5	8.30	−2.513	0.572	−1.288	158.644
35	4−CH_3,  6−	COOH	4	8.39	−2.068	0.655	−0.949	146.895
36	4−CH_3,  6−	Tetrazol−5−yl	3	8.52	−1.744	0.661	−1.234	157.531
37	4−CH_3,  6−	Tetrazol−5−yl	5	8.30	−2.229	0.859	−1.056	148.099
38	 6−	Tetrazol−5−yl	11	7.95	−1.729	0.818	−0.963	143.058
39	 6−	Tetrazol−5−yl	240	6.61	−2.117	0.494	−1.310	143.981
40	4−CH_3, 6-$(CH_3)_2NCH_2$	COOH	158	6.80	−2.427	0.696	−0.335	133.398

[a] IC_{50} for specific binding of [^{125}I]AII to rat lung membrane preparation; [b] logarithmic value of IC_{50}; [c] MORSE descriptor; [d] Molar refractivity

statistics (evaluated as the ratio of the equivalent regression equations to the total number of randomized sets; a chance value of 0.001 corresponds to 0.1% chance of fortuitous correlation), outliers (on the basis of Z-score value) and predictive squared correlation coefficient of test set (r^2_{pred}).

Results and Discussion

In the present study, an attempt has been made to find structural requirement for the antagonistic activity of 6−substituted benzimidazole analogs as AII AT1 receptor using the Hansch approach. The multivariant expressions were developed on the basis of adjustable correlation coefficient (r^2_{adj}). This parameter tells us the statistical significance of incorporated physicochemical descriptor in SEQ−MLR. Adjustable r^2 takes into account the adjustment of conventional correlation coefficient (r^2). Therefore, if a physicochemical descriptor is added that does not contribute its fair share, the r^2_{adj} will actually decline. Adjustable correlation coefficient is a measure of the percentage explained variation in the dependent variable that takes into account the relationship between the number of cases and the number of independent variable in the regression model. Whereas r^2 will always increase when an independent variable is added, r^2_{adj} will decrease if the added variable does reduce the unexplained variation enough to offset the loss of degrees of freedom.

It is necessary that the proposed models should have both the statistical quality as well as good predictive power therefore all the expressions were tested for internal and external validation. Both the validation put forward decision making input for selection of QSAR models. Internal validation was carried out using leave one out (LOO) cross validation method, bootstrapping technique and randomized biological activity test while external validation confirmed with test set data. Tri−variant expressions (Eq. 1 & 2) which fulfill all the validation criteria up to significant echelon were considered as QSAR model−1 and 2 respectively.

Eq. 1. pIC_{50} = Mor21u [1.418 (±0.682)] + Mor31u [1.945 (±0.635)] + Mor14m [1.277 (±0.618)] + MR [0.083 (±0.0163)] − 1.53081 (±1.75312)

n=28, r=0.919, r^2=0.844, SE=0.419, F=31.114

Eq. 2. pIC_{50} = Mor13u [0.854 (± 0.389)] − Mor18m [2.025 (±0.702)] − Mor28m [3.113 (± 1.658)] − Mor26v [−2.511 (±1.123)] + 2.399 (±1.122)

n=28, r=0.918, r^2=0.843, SE=0.421, F=30.834

Tab. 2. QSAR statistics of significant models

Model No.	r^{2a}	r^2_{adj} [b]	SEE[c]	F[d]	r^2_{bs} [e]	S_{bs} [f]	Chance	Q^{2g}	S_{PRESS} [h]	S_{DEP} [i]	r^2_{pred} [j]	Out−liers
1	0.84	0.82	0.41	31.11	0.85	0.058	0.001	0.78	0.40	0.40	0.40	Nil
2	0.84	0.82	0.42	30.83	0.86	0.053	0.001	0.70	0.57	0.52	0.15	Nil

[a] squared correlation coefficient; [b] adjustable correlation coefficient; [c] standard error of estimate; [d] sequential Fischer test; [e] bootstrapping squared correlation coefficient; [f] bootstrapping standard deviation; [g] cross validated squared correlation coefficient using leave one out procedure; [h] squared sum of predicted residual; [i] standard error of prediction; [j] predictive squared correlation coefficient of test set.

Both the models have correlation coefficient more than (r=0.900), which accounts for more than 84% of the variance in the activity. The inter−correlation among the parameters used in model no.1 is less than 0.370 (Table 2 & 3).

Tab. 3. Inter−correlation matrix of descriptors used in model 1

Descriptors	**Mor21u**	**Mor31u**	**Mor14m**	**MR**
Mor21u	1.000			
Mor31u	0.285	1.000		
Mor14m	0.157	0.294	1.000	
MR	0.370	0.242	0.258	1.000

The model shows, that in a multi−variant model, a dependent variable can be predicted from a linear combination of the independent variables. The P value is less than 0.01 for each physiochemical parameter involved in model generation. The data showed overall internal statistical significance level better than 99.9% as it exceeded from the tabulated $F_{(4,25\ \alpha\ 0.001)}$ = 10.8. Models were further tested for outliers by Z−score method and no compound was found to be an outlier (Table 2) which suggested that the models are able to explain the structurally diverse analogs that is helpful in designing of more potent compounds using physiochemical parameters.

The Z−value for individual compounds within the specific range (<j2.5 j) indicated absence of outliers.

It is used to label the outliers in the dataset.

$$z_i = \frac{(x_i - \bar{x})}{s}$$

where

$$s = \sqrt{\frac{\sum_{i-1}^{n}(x_i - \bar{x})^2}{n-1}}$$

Where,

n = Number of datapoints
x_i = estimated value of predicted
$\bar{x}$ = Mean of observed values of predicted

Only high correlation coefficient is not enough to select the equation as a model and hence various statistical approaches were used to confirm the robustness and practical applicability of the equations. Model 1 & 2 showed that their probability of chance correlation is less than 0.1% in randomized biological activity test. Bootstrapping technique was employed to confirm the contribution of descriptors to the activity. They were equi−intense or of different rank. The value of bootstrapping squared correlation coefficient and bootstrapping standard deviation implies that the equations are proper representative of the group of analogs (Table 2).

The internal consistency of the training set was confirmed by using leave one out (LOO) cross validation method to ensure the robustness of the Model. Models showed good internal consistency (Q^2 = 0.70 & 0.78 see table 2), which reduces the probability of coincidental correlation of the expression (Table 4, Figure1). Although model showed good internal consistency, they may not be applicable for the analogs which were never used in the generation of the correlation. Therefore, the external extrapolation power of the model was further authenticated by a test set of twelve compounds. Values of predictive squared correlation coefficient ($r^2_{pred)}$ are 0.40 and 0.15 for model. 1 & 2 respectively (see Table 2)

Tab. 4. Calculated and Predicted pIC_{50} (by LOO method) of training set with residual and Z−score value using model 1

Comp No.	Model				
	Calculated pIC_{50}	Residual	Z−value	Predicted pIC_{50} (loo)	Residual (loo)
T−1	5.79	0.60	1.56	5.68	0.71
T−2	6.60	−0.68	−1.77	6.70	−0.78
T−3	5.91	0.01	0.01	5.91	0.01
T−4	5.91	0.16	0.40	5.89	0.18
T−5	6.35	−0.04	−0.09	6.36	−0.05
T−6	5.58	0.18	0.46	5.55	0.21
T−7	6.21	−0.13	−0.33	6.24	−0.16
T−8	5.90	0.36	0.92	5.84	0.42
T−10	5.96	−0.72	−1.87	6.07	−0.83
T−11	6.46	−0.13	0.32	6.47	−0.14
T−15	6.58	−0.49	−1.26	6.79	−0.70
T−17	7.41	−0.62	−1.61	7.51	−0.72
T−19	6.82	−0.03	−0.06	6.82	−0.03
T−20	7.25	−0.19	0.49	7.27	−0.21
T−22	7.08	0.50	1.29	6.98	0.60
T−23	7.89	−0.41	−1.07	7.94	−0.46
T−24	6.90	0.56	1.46	6.86	0.6
T−26	7.63	0.04	0.12	7.20	0.47
T−28	7.51	0.58	1.49	7.46	0.63
T−30	8.45	−0.06	0.14	6.48	1.91
T−31	8.36	0.16	0.41	8.33	0.19
T−32	8.57	−0.05	0.14	8.61	−0.09
T−33	7.95	−0.07	−0.18	7.97	−0.09
T−35	7.77	0.62	1.61	7.72	0.67
T−37	7.90	0.40	1.02	7.85	0.45
T−38	8.23	−0.28	−0.71	8.30	−0.35
T−39	6.68	−0.07	0.17	6.70	−0.09
T−40	7.00	−0.20	−0.53	7.03	−0.23

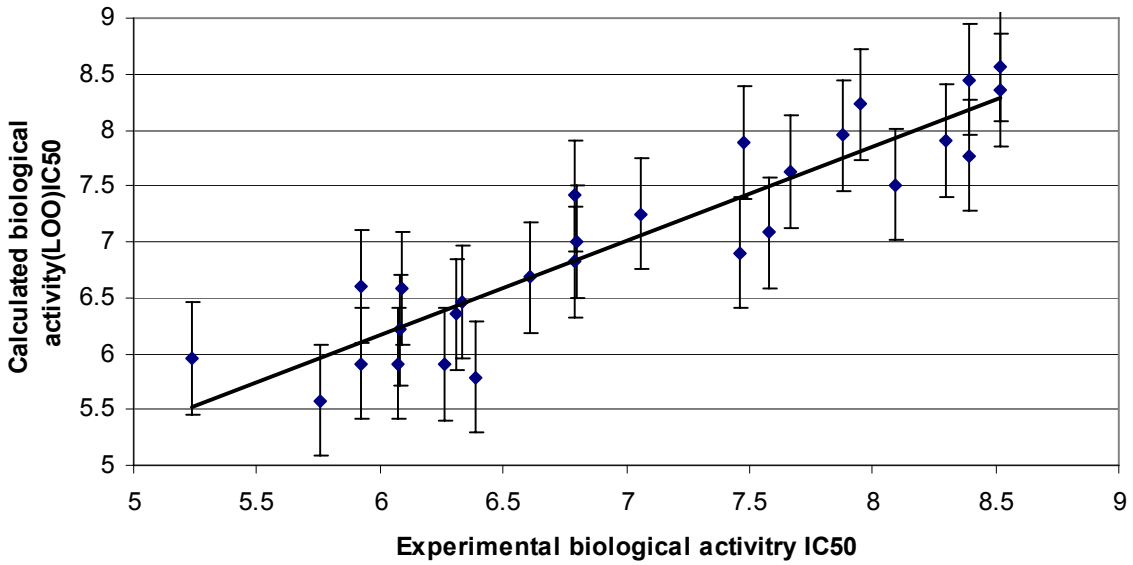


Fig. 1. A plot between observed plC_{50} and calculated (LOO) plC_{50} with residual using model 1

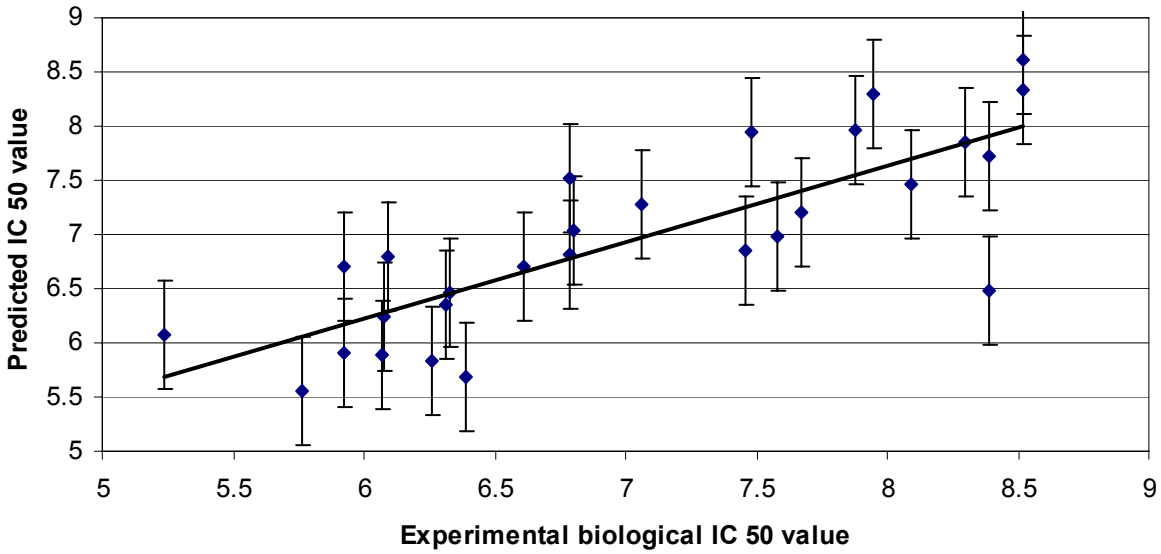


Fig. 2. A plot between observed plC_{50} and predicted (LOO) plC_{50} with residual using model 1

The test set supported significantly the robustness, predictiveness and wide applicability of the model 1 over model 2 (Table 5 & Figure 2). In general the model fulfills the statistical validation criteria to the significant extent.

Tab. 5. Observed and Predicted pIC_{50} of test set with residual using model 2

S. No.	Observed pIC_{50}	Predicted pIC_{50}	Residual
Test−9	5.97	6.29	−0.32
Test−12	6.74	6.82	−0.08
Test−13	5.74	5.90	−0.16
Test−14	4.53	6.25	−1.72
Test−16	7.58	6.63	0.95
Test−18	6.40	7.57	−1.17
Test−21	8.52	5.65	2.87
Test−25	7.09	7.15	−0.06
Test−27	8.00	7.77	0.23
Test−29	8.52	8.17	0.35
Test−34	8.30	7.51	0.79
Test−36	8.52	8.76	−0.24

Mor 21u, Mor31u and Mor14m are 3D molecular representation of structure based on electron diffraction code (MoRSE code) contributing positively to the biological activity. [10−12] was calculated by summing atom weights viewed by a different angular scattering function. The values of these code functions were calculated at 32 evenly distributed values of scattering angle(s) in the range of 0−31 $Å^{-1}$ from the three dimensional atomic co−ordinates of a molecule. The 3D−MoRSE code was calculated using following expression;

$$I(s) = \sum_{i=2}^{N} \sum_{j=1}^{i-1} A_i \cdot A_j \cdot \frac{\sin(sr_{ij})}{sr_{ij}}$$

Where,
s is scattering angle
r_{ij} is interatomic distance of i^{th} and j^{th} atom
A_i and A_j are atomic properties of i^{th} and j^{th} atom respectively including vander Waals volume, atomic number, atomic mass, partial atomic charges, residual electro−negativities, and atom polarizability.

Model 1 indicates that Molar refractivity [13] is contributing positively to the biological activity showing that the steric interaction play role in receptor binding.

From the discussion it concludes that three dimensional structural properties (MoRSE codes) and molar refractivity of 6−substituted benzimidazole derivative is decisive in AT1 receptor AII antagonist activity.

Acknowledgment

The authors are thankful to the HOD, School of pharmacy, DAVV, Indore for providing facilities to complete this work.

Authors' Statement

Competing interest

The authors declare no conflict of interest.

References

[1] Gilman G, Goodman SL, Rall TW, Murad F.
In: The Pharmacological Basis of Therapeutics.7th ed.
New York: Macmillan Publishing Co, 1985: 639–659.

[2] Smith RD, Chiu AT, Wong PC, Herblin WF, Timmermans PBMWM.
Pharmacology of Nonpeptide Angiotensin II Receptor Antagonist.
Ann Rev Pharmacol Toxicol. 1992; 32: 135–165.
doi:10.1146/annurev.pa.32.040192.001031

[3] Pandya T, Pandey SK, Tiwari M, Chaturvedi SC, Saxena AK.
3D−QSAR studies of triazolinone based balanced AT_1/AT_2 receptor antagonist.
Bioorg Med Chem. 2001; 9: 291–300.
doi:10.1016/S0968−0896(00)00243−1

[4] Pandya T, Chaturvedi SC.
Structure−activity relationship study of some triazolinone based compounds with antagonistic balanced activity on angiotensin II receptor subtypes AT1 and AT2. A three−dimensional quantitative structure−activity relationship investigation.
Arzneimittelforschung. 2005; 55: 265–270.
PMid:15960425

[5] Jain A., Chaturvedi SC.
Rationalization of Physicochemical property of some substituted benzimidazole bearing acidic heterocyclic towards angiotensin II antagonist: a QSAR approach.
Asian J Biochem. 2008; 3: 330–336.
doi:10.3923/ajb.2008.330.336

[6] Uwe JR, Gerhard M, Berthold N, Kain MH, Helmut M, Michael E, Jecobus CA, Vanmeel W, Nobert HH.
6−Substituted benzimidazoles as new nonpeptide angiotensin II receptor antagonists: synthesis, biological activity, and structure−activity relationships.
J Med Chem. 1993; 36: 4040–4051.
doi:10.1021/jm00077a007

[7] CS Chem Office, version 8.0.
Cambridge soft Corporation, software publishers association,1730 m Street, suite 700, Washington DC 20036, USA.

[8] Todeschini R, Consonni V.
DRAGON−Software for the calculation of molecular Descriptors.
2001; rel. 1.12 for Windows.

[9] Gupta AK, Arokia BM, Khakhedikar SG.
Valstat: validated program for Quantitative structure activity relation ship studies.
Indian J Pharm Sci. 2004; 66: 396–402.

[10] Gasteiger J, Sadowski J, Schuur J, Selzer P, Steinhauer L, Steinhauer V.
Chemical information in 3D space.
J Chem Inf Comput Sci. 1996; 36: 1030–1037.
doi:10.1021/ci960343+

[11] Schuur J, Selzer HP, Gasteiger J.
The coding of the three –dimensional structure of molecules by molecule transform and its application structure−spectra correlation and studies of biological activity.
J Chem Inf Comput Sci. 1996; 36: 334–344.
doi:10.1021/ci950164c

[12] Schuur J, Gasteiger J.
Infrared spectra simulation of substituted benzene derivatives on the basis of a 3D structure representation.
Anal Chem. 1997; 83: 2398–2405.
doi:10.1021/ac9611071

[13] Viswanadhan VN,Ghose AK, Ravankar GR, Robins RK.
Atomic physicochemical parameters for three dimensional structure directed quantitative structure−activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics.
J Chem Inf Comput Sci. 1989; 29: 163–172.
doi:10.1021/ci00063a006